World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:15, No:01, 2021

End-to-End Spanish-English Sequence Learning Translation Model

Authors : Vidhu Mitha Goutham, Ruma Mukherjee

Abstract : The low availability of well-trained, unlimited, dynamic-access models for specific languages makes it hard for corporate users to adopt quick translation techniques and incorporate them into product solutions. As translation tasks increasingly require a dynamic sequence learning curve; stable, cost-free opensource models are scarce. We survey and compare current translation techniques and propose a modified sequence to sequence model repurposed with attention techniques. Sequence learning using an encoder-decoder model is now paving the path for higher precision levels in translation. Using a Convolutional Neural Network (CNN) encoder and a Recurrent Neural Network (RNN) decoder background, we use Fairseq tools to produce an end-to-end bilingually trained Spanish-English machine translation model including source language detection. We acquire competitive results using a duo-lingo-corpus trained model to provide for prospective, ready-made plug-in use for compound sentences and document translations. Our model serves a decent system for large, organizational data translation needs. While acknowledging its shortcomings and future scope, it also identifies itself as a well-optimized deep neural network model and solution.

Keywords : attention, encoder-decoder, Fairseq, Seq2Seq, Spanish, translation
Conference Title : ICDMBDDDM 2021 : International Conference on Data Mining, Big Data, Database and Data Management
Conference Location : New York, United States
Conference Dates : January 28-29, 2021